* **Rows before preprocessing:** 257,673
* **Classes:**  
  *['Analysis', 'Backdoor', 'DoS', 'Exploits', 'Fuzzers', 'Generic', 'Normal', 'Reconnaissance', 'Shellcode', 'Worms']*
* **Rows after preprocessing:** 257,673
  + *Number of features:* 41
* **Global Split:**
  + **Training set:** 206,138 samples
  + **Test set:** 51,535 samples

**🧬 Data Counts Per Class**

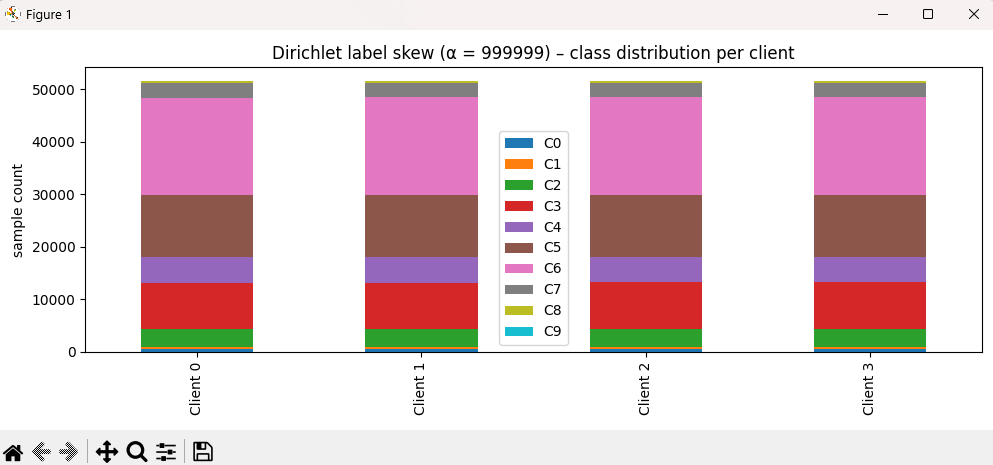
| **Class ID** | **Class Name** | **Train Count** | **Test Count** |
| --- | --- | --- | --- |
| 0 | Analysis | 2,142 | 535 |
| 1 | Backdoor | 1,863 | 466 |
| 2 | DoS | 13,082 | 3,271 |
| 3 | Exploits | 35,620 | 8,905 |
| 4 | Fuzzers | 19,397 | 4,849 |
| 5 | Generic | 47,097 | 11,774 |
| 6 | Normal | 74,400 | 18,600 |
| 7 | Reconnaissance | 11,189 | 2,798 |
| 8 | Shellcode | 1,209 | 302 |
| 9 | Worms | 139 | 35 |

* **Total Features:** 41
* **Total Classes:** 10

**Local epochs = 3, Rounds = 50.**

**RUN 1:**

IID-data, no DP



Duration: 152.02 seconds

**Overall Performance (Round 50)**  
• Global Accuracy: 86.8%  
• Global F1 Score (macro): 0.573  
• Global Precision (macro): 0.739  
• Global Recall (macro): 0.545

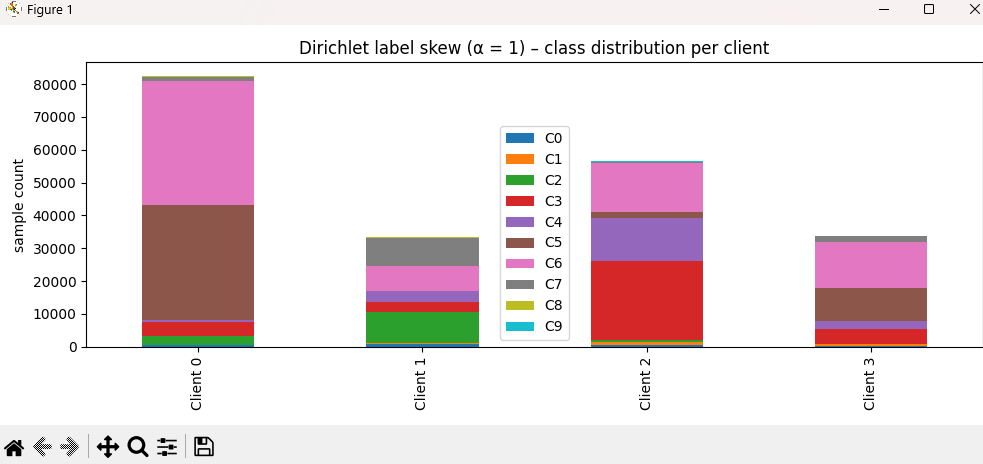
**Global Loss Trends**  
• Distributed Training Loss: decreased from 0.530 → 0.343  
• Centralized Evaluation Loss: decreased from 2.313 → 0.343

**Per‑Class Metrics at Round 50**

| **Class** | **F1** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.276 | 0.802 | 0.166 | Precision strong, recall low |
| 1 | 0.057 | 0.667 | 0.030 | Poor detection, slight gain |
| 2 | 0.287 | 0.411 | 0.221 | Moderate recall, needs more |
| 3 | 0.726 | 0.638 | 0.841 | Balanced and robust |
| 4 | 0.835 | 0.841 | 0.828 | Very strong across metrics |
| 5 | 0.983 | 0.994 | 0.973 | Near-perfect performance |
| 6 | ~1.000 | ~1.000 | ~1.000 | Perfect classification |
| 7 | 0.792 | 0.824 | 0.764 | Well-balanced |
| 8 | 0.462 | 0.535 | 0.407 | Moderate performance, much improved |
| 9 | 0.311 | 0.700 | 0.200 | Same detection as non‑IID run |

**Run 2:**

Non-IID alpha = 1

****

Duration: 139.68 seconds  
Overall Performance (Round 50)  
• Global Accuracy: 86.5%  
• Global F1 Score (macro): 0.526  
• Global Precision (macro): 0.764  
• Global Recall (macro): 0.524

Global Loss Trends  
• Distributed Training Loss: 0.631 → 0.363  
• Centralized Evaluation Loss: 2.27 → 0.363

Per-Class Metrics at Round 50

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.278 | 0.689 | 0.174 | Low recall, decent precision |
| 1 | 0.017 | 0.667 | 0.009 | Model still failing this class |
| 2 | 0.020 | 0.472 | 0.010 | Slight improvement, still weak |
| 3 | 0.728 | 0.590 | 0.948 | High recall, solid precision |
| 4 | 0.843 | 0.894 | 0.797 | Very strong across all metrics |
| 5 | 0.982 | 0.994 | 0.970 | Near-perfect classification |
| 6 | 1.000 | 1.000 | 1.000 | Perfect classification |
| 7 | 0.776 | 0.878 | 0.695 | Strong but slightly recall-limited |
| 8 | 0.512 | 0.459 | 0.579 | Notable improvement from prior runs |
| 9 | 0.108 | 1.000 | 0.057 | Rare correct predictions |

**Run 3:**

Same as before.

Alpha = 0.7

**Duration:** 118.85 seconds  
**Overall Performance (Round 50)**  
• Global Accuracy: 86.7%  
• Global F1 Score (macro): 0.522  
• Global Precision (macro): 0.688  
• Global Recall (macro): 0.508

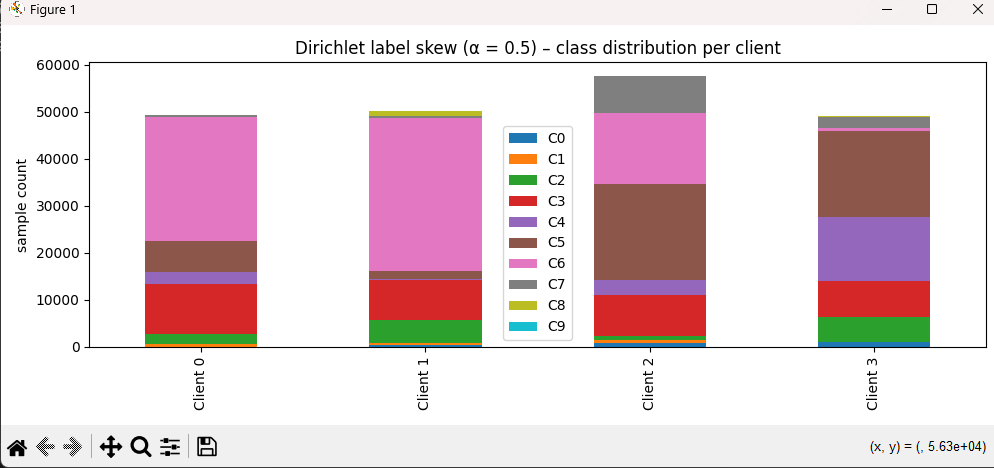
**Global Loss Trends**  
• Distributed Training Loss: 0.601 → 0.356  
• Centralized Evaluation Loss: 2.318 → 0.352

**Per-Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.211 | 0.901 | 0.120 | Low recall, high precision |
| 1 | 0.042 | 0.667 | 0.021 | Almost no learning occurred |
| 2 | 0.143 | 0.414 | 0.086 | Weak detection |
| 3 | 0.728 | 0.603 | 0.918 | High recall, consistent |
| 4 | 0.844 | 0.879 | 0.812 | Strong and stable |
| 5 | 0.982 | 0.995 | 0.969 | Near-perfect learning |
| 6 | 1.000 | 1.000 | 1.000 | Perfect classification |
| 7 | 0.791 | 0.849 | 0.740 | Balanced and robust |
| 8 | 0.482 | 0.576 | 0.414 | Improved over prior runs |
| 9 | 0.000 | 0.000 | 0.000 | Completely unlearned |

**Run 4:**

Non-IID alpha = 0.5



Duration: **146.60 seconds**

**Overall Performance (Round 50)**

* Global Accuracy: 86.2%
* Global F1 Score (macro): 0.497
* Global Precision (macro): 0.672
* Global Recall (macro): 0.481

**Global Loss Trends**

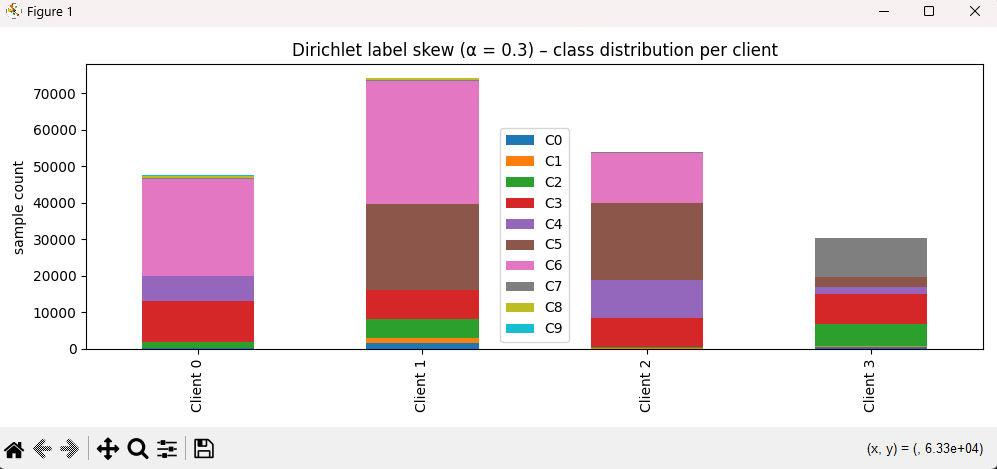
* Distributed Training Loss: 0.625 → 0.364
* Centralized Evaluation Loss: 2.35 → 0.362

**Per-Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.222 | 0.872 | 0.127 | Low recall, high precision |
| 1 | 0.000 | 0.000 | 0.000 | Model completely failed |
| 2 | 0.086 | 0.405 | 0.048 | Barely learned this class |
| 3 | 0.722 | 0.584 | 0.946 | High recall, moderate F1 |
| 4 | 0.826 | 0.932 | 0.742 | Strong performance |
| 5 | 0.982 | 0.994 | 0.971 | Near-perfect learning |
| 6 | 1.000 | 1.000 | 1.000 | Perfect classification |
| 7 | 0.782 | 0.814 | 0.751 | Balanced and strong |
| 8 | 0.297 | 0.621 | 0.195 | Weak detection |
| 9 | 0.054 | 0.500 | 0.029 | Very late and poor detection |

**Run 5:**

Non-IID alpha = 0.3



**Duration:** 160.38 seconds  
**Overall Performance (Round 50)**  
• Global Accuracy: 83.6%  
• Global F1 Score (macro): 0.470  
• Global Precision (macro): 0.708  
• Global Recall (macro): 0.446

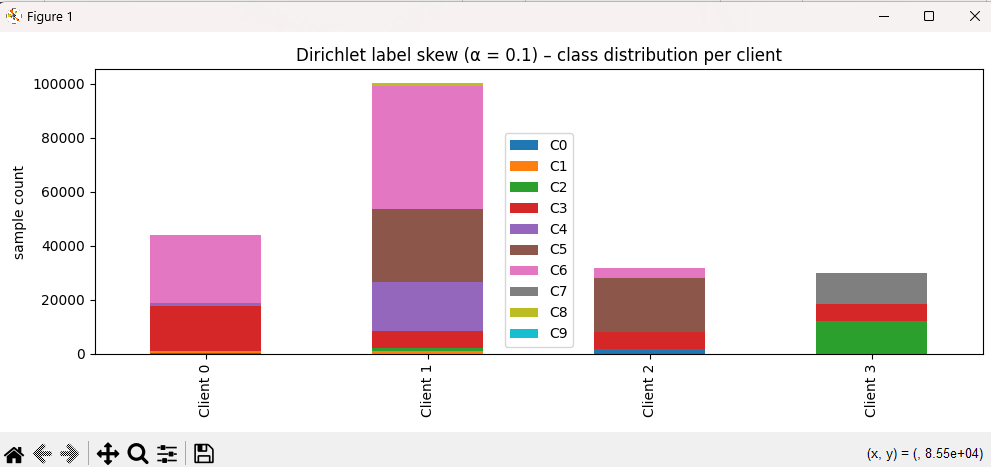
**Global Loss Trends**  
• Distributed Training Loss: 0.818 → 0.468  
• Centralized Evaluation Loss: 2.32 → 0.470

**Per-Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.232 | 0.828 | 0.135 | Low recall, high precision |
| 1 | 0.000 | 0.000 | 0.000 | Model completely failed |
| 2 | 0.052 | 0.508 | 0.028 | Barely learned this class |
| 3 | 0.686 | 0.526 | 0.986 | Very high recall |
| 4 | 0.640 | 0.963 | 0.480 | Strong precision, recall lag |
| 5 | 0.981 | 0.991 | 0.971 | Near-perfect classification |
| 6 | 1.000 | 1.000 | 1.000 | Perfect classification |
| 7 | 0.745 | 0.898 | 0.637 | Balanced and robust |
| 8 | 0.051 | 0.667 | 0.026 | Very low recall |
| 9 | 0.311 | 0.700 | 0.200 | Late but moderate detection |

**Run 6:**

Non-IID alpha = 0.1



**Duration:** 171.48 seconds  
**Overall Performance (Round 50)**  
• Global Accuracy: 81.9%  
• Global F1 Score (macro): 0.371  
• Global Precision (macro): 0.735  
• Global Recall (macro): 0.381

**Global Loss Trends**  
• Distributed Training Loss: 0.92 → 0.72  
• Centralized Evaluation Loss: 2.36 → 0.72

**Per-Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.040 | 1.000 | 0.021 | Barely learned, but very high precision |
| 1 | 0.004 | 1.000 | 0.002 | Completely failed again |
| 2 | 0.002 | 1.000 | 0.001 | No significant improvement |
| 3 | 0.659 | 0.495 | 0.985 | High recall, moderate precision |
| 4 | 0.792 | 0.917 | 0.698 | Best balance among all classes |
| 5 | 0.983 | 0.995 | 0.971 | Near-perfect classification |
| 6 | 1.000 | 1.000 | 1.000 | Perfect classification |
| 7 | 0.000 | 0.000 | 0.000 | Not learned at all |
| 8 | 0.037 | 0.273 | 0.020 | Low precision and recall |
| 9 | 0.195 | 0.667 | 0.114 | Detectable but weak |

With 10 clients:

**Run 7:**

IID

**Duration:** 82.99 seconds  
**Rounds:** 50

**Overall Performance (Round 50)**  
• Global Accuracy: 86.33%  
• Global F1 Score (macro): 0.5297  
• Global Precision (macro): 0.6948  
• Global Recall (macro): 0.5073

**Global Loss Trends**  
• Distributed Training Loss: 0.637 → 0.360 (round 1→50)  
• Centralized Evaluation Loss: 2.34 → 0.362

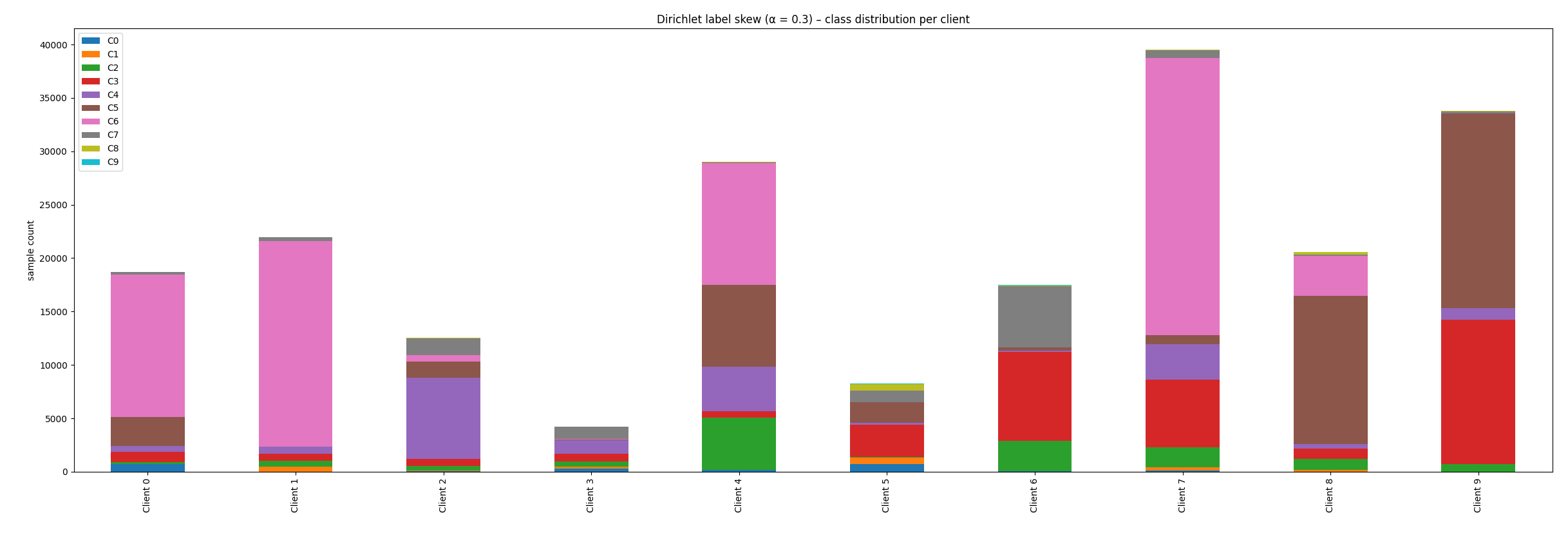
**Per Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.260 | 0.757 | 0.157 | Slightly improved detection of minority class |
| 1 | 0.000 | 0.000 | 0.000 | Remains undetectable |
| 2 | 0.270 | 0.383 | 0.209 | Gradual improvement, still low recall |
| 3 | 0.722 | 0.630 | 0.844 | Consistent and strong |
| 4 | 0.823 | 0.822 | 0.823 | Balanced and high-performing |
| 5 | 0.981 | 0.992 | 0.969 | Near-perfect performance |
| 6 | 1.000 | 1.000 | 1.000 | Perfect classification |
| 7 | 0.785 | 0.813 | 0.758 | Reliable and steady |
| 8 | 0.348 | 0.550 | 0.255 | Minor class, improved significantly |
| 9 | 0.108 | 1.000 | 0.057 | Very low recall despite perfect precision |

**Run 8:**

**10 CLIENTS**

Non-IID alpha = 0.3



**Duration:** 94.49 seconds  
**Rounds:** 50

**Overall Performance (Round 50)**  
• Global Accuracy: 85.92%  
• Global F1 Score (macro): 0.4991  
• Global Precision (macro): 0.6157  
• Global Recall (macro): 0.4847

**Global Loss Trends**  
• Distributed Training Loss: 1.03 → 0.38 (round 1→50)  
• Centralized Evaluation Loss: 2.29 → 0.39

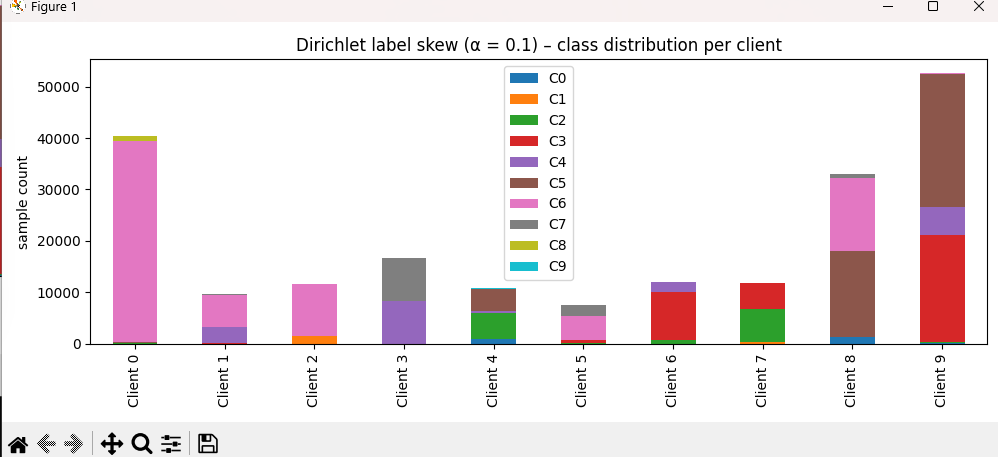
**Per Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.125 | 0.878 | 0.067 | Slight recovery, still low recall despite good precision |
| 1 | 0.000 | 0.000 | 0.000 | Undetected — model fails to recognize this class entirely |
| 2 | 0.399 | 0.387 | 0.412 | Stable but modest performance; precision slightly improved |
| 3 | 0.711 | 0.645 | 0.791 | Strong recall; consistent precision |
| 4 | 0.821 | 0.827 | 0.815 | High, balanced classification |
| 5 | 0.982 | 0.996 | 0.968 | Near-perfect, as expected |
| 6 | 1.000 | 1.000 | 1.000 | Perfect classification |
| 7 | 0.747 | 0.848 | 0.668 | Good precision, improving recall |
| 8 | 0.207 | 0.576 | 0.126 | Significant gains, but still weak in recall |
| 9 | 0.000 | 0.000 | 0.000 | Undetected — class not captured at all |

**Run 9:**

**10 CLIENTS**

Non-IID alpha = 0.1



**Duration:** 119.08 seconds  
**Rounds:** 50

**Overall Performance (Round 50)**  
• Global Accuracy: 82.69%  
• Global F1 Score (macro): 0.4929  
• Global Precision (macro): 0.5751  
• Global Recall (macro): 0.4956

**Global Loss Trends**  
• Distributed Training Loss: 1.31 → 0.47 (round 1→50)  
• Centralized Evaluation Loss: 2.36 → 0.47

**Per Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.1980 | 0.8451 | 0.1121 | Very low recall despite strong precision |
| 1 | 0.0000 | 0.0000 | 0.0000 | Completely undetected throughout training |
| 2 | 0.4538 | 0.3385 | 0.6879 | Strong recall, modest precision |
| 3 | 0.6419 | 0.6922 | 0.5984 | Balanced, consistent class |
| 4 | 0.7860 | 0.8296 | 0.7468 | High performance overall |
| 5 | 0.9696 | 0.9616 | 0.9777 | Near-perfect detection |
| 6 | 0.9997 | 0.9993 | 1.000 | Perfect classification |
| 7 | 0.5339 | 0.7567 | 0.4124 | Improved late training; solid precision |
| 8 | 0.1826 | 0.1359 | 0.2781 | Detection improved, but still weak |
| 9 | 0.1639 | 0.1923 | 0.1429 | Slight late detection, low frequency |

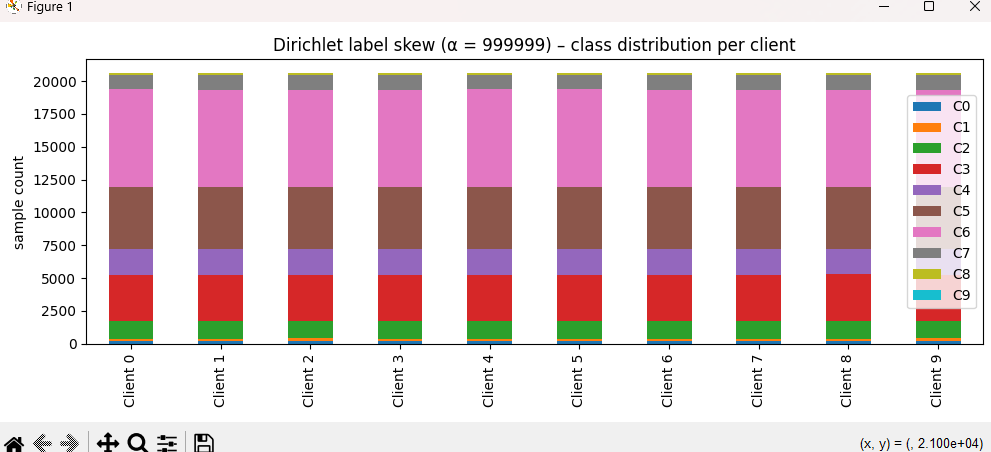
**Run 10:**

**FRACTION FIT = 0.5**

**FRACTION EVAL = 1**

**10 CLIENTS**

**IID**



**Duration:** 75.86 seconds  
**Rounds:** 50

**Overall Performance (Round 50)**  
• Global Accuracy: 86.29%  
• Global F1 Score (macro): 0.5280  
• Global Precision (macro): 0.6667  
• Global Recall (macro): 0.5087

**Global Loss Trends**  
• Distributed Training Loss: 0.621 → 0.362 (round 1→50)  
• Centralized Evaluation Loss: 2.35 → 0.364

**Per Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.250 | 0.769 | 0.150 | Slight improvement; precision high, recall still poor |
| 1 | 0.000 | 0.000 | 0.000 | Undetected across all rounds |
| 2 | 0.204 | 0.410 | 0.135 | Better than before, but recall weak |
| 3 | 0.722 | 0.617 | 0.870 | Strong class, high recall consistency |
| 4 | 0.822 | 0.824 | 0.820 | High stability, balanced precision and recall |
| 5 | 0.981 | 0.994 | 0.969 | Near-perfect, stable across rounds |
| 6 | 1.000 | 0.9999 | 1.000 | Perfect from early rounds |
| 7 | 0.767 | 0.781 | 0.754 | Consistently strong performance |
| 8 | 0.428 | 0.606 | 0.331 | Noticeable improvement in minority class |
| 9 | 0.105 | 0.667 | 0.057 | Better precision, detection remains minimal |

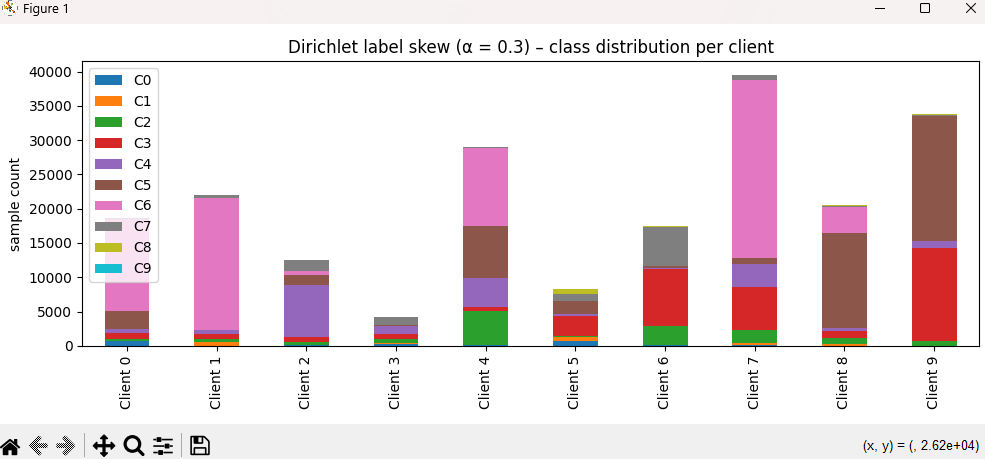
**Run 11:**

**FRACTION FIT = 0.5**

**FRACTION EVAL = 1**

**10 CLIENTS**

Non-IID alpha = 0.3



**Duration:** 100.74 seconds  
**Rounds:** 50

**Overall Performance (Round 50)**  
• Global Accuracy: 85.88%  
• Global F1 Score (macro): 0.4908  
• Global Precision (macro): 0.6146  
• Global Recall (macro): 0.4781

**Global Loss Trends**  
• Distributed Training Loss: 0.958 → 0.385 (round 1→50)  
• Centralized Evaluation Loss: 2.39 → 0.387

**Per Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.0854 | 0.8889 | 0.0449 | Precision high, but recall very low |
| 1 | 0.0000 | 0.0000 | 0.0000 | Completely undetected |
| 2 | 0.3396 | 0.3845 | 0.3042 | Moderate performance, low recall |
| 3 | 0.7135 | 0.6341 | 0.8158 | High recall, solid overall |
| 4 | 0.8113 | 0.7906 | 0.8332 | Very strong and balanced |
| 5 | 0.9814 | 0.9954 | 0.9677 | Near-perfect |
| 6 | 1.0000 | 1.0000 | 1.0000 | Perfect |
| 7 | 0.7547 | 0.8494 | 0.6791 | Improved; high precision |
| 8 | 0.2216 | 0.6029 | 0.1358 | Precision much higher than recall |
| 9 | 0.0000 | 0.0000 | 0.0000 | Completely undetected |

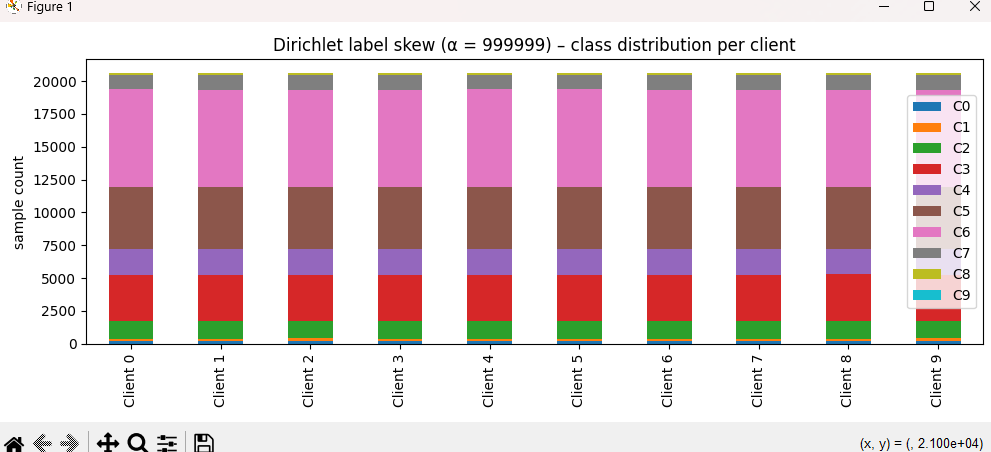
**Run 12:**

**FRACTION FIT = 0.5**

**FRACTION EVAL = 1**

**10 CLIENTS**

**IID**



**Duration:** 75.86 seconds  
**Rounds:** 50

**Overall Performance (Round 50)**  
• Global Accuracy: 86.29%  
• Global F1 Score (macro): 0.5280  
• Global Precision (macro): 0.6667  
• Global Recall (macro): 0.5087

**Global Loss Trends**  
• Distributed Training Loss: 0.621 → 0.362 (round 1→50)  
• Centralized Evaluation Loss: 2.35 → 0.364

**Per Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.250 | 0.769 | 0.150 | Slight improvement; precision high, recall still poor |
| 1 | 0.000 | 0.000 | 0.000 | Undetected across all rounds |
| 2 | 0.204 | 0.410 | 0.135 | Better than before, but recall weak |
| 3 | 0.722 | 0.617 | 0.870 | Strong class, high recall consistency |
| 4 | 0.822 | 0.824 | 0.820 | High stability, balanced precision and recall |
| 5 | 0.981 | 0.994 | 0.969 | Near-perfect, stable across rounds |
| 6 | 1.000 | 0.9999 | 1.000 | Perfect from early rounds |
| 7 | 0.767 | 0.781 | 0.754 | Consistently strong performance |
| 8 | 0.428 | 0.606 | 0.331 | Noticeable improvement in minority class |
| 9 | 0.105 | 0.667 | 0.057 | Better precision, detection remains minimal |

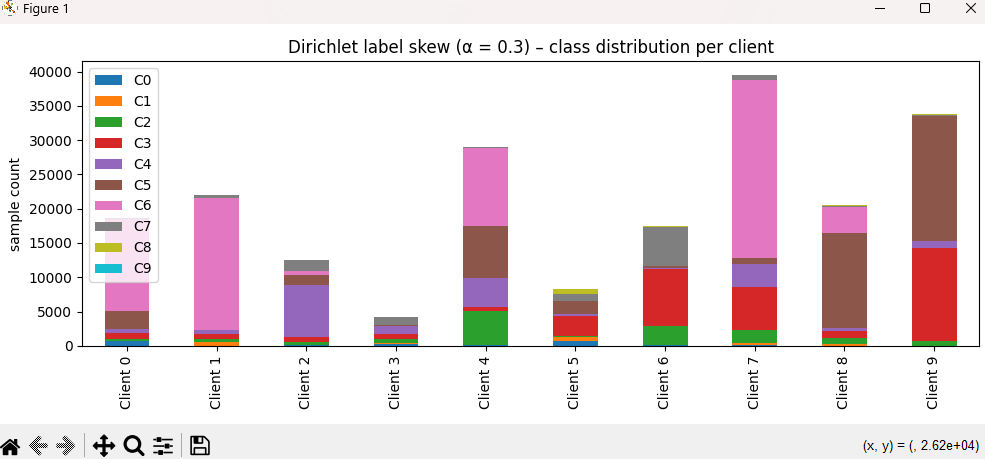
**Run 13:**

**FRACTION FIT = 0.5**

**FRACTION EVAL = 1**

**10 CLIENTS**

Non-IID alpha = 0.3



**Duration:** 100.74 seconds  
**Rounds:** 50

**Overall Performance (Round 50)**  
• Global Accuracy: 85.88%  
• Global F1 Score (macro): 0.4908  
• Global Precision (macro): 0.6146  
• Global Recall (macro): 0.4781

**Global Loss Trends**  
• Distributed Training Loss: 0.958 → 0.385 (round 1→50)  
• Centralized Evaluation Loss: 2.39 → 0.387

**Per Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.0854 | 0.8889 | 0.0449 | Precision high, but recall very low |
| 1 | 0.0000 | 0.0000 | 0.0000 | Completely undetected |
| 2 | 0.3396 | 0.3845 | 0.3042 | Moderate performance, low recall |
| 3 | 0.7135 | 0.6341 | 0.8158 | High recall, solid overall |
| 4 | 0.8113 | 0.7906 | 0.8332 | Very strong and balanced |
| 5 | 0.9814 | 0.9954 | 0.9677 | Near-perfect |
| 6 | 1.0000 | 1.0000 | 1.0000 | Perfect |
| 7 | 0.7547 | 0.8494 | 0.6791 | Improved; high precision |
| 8 | 0.2216 | 0.6029 | 0.1358 | Precision much higher than recall |
| 9 | 0.0000 | 0.0000 | 0.0000 | Completely undetected |

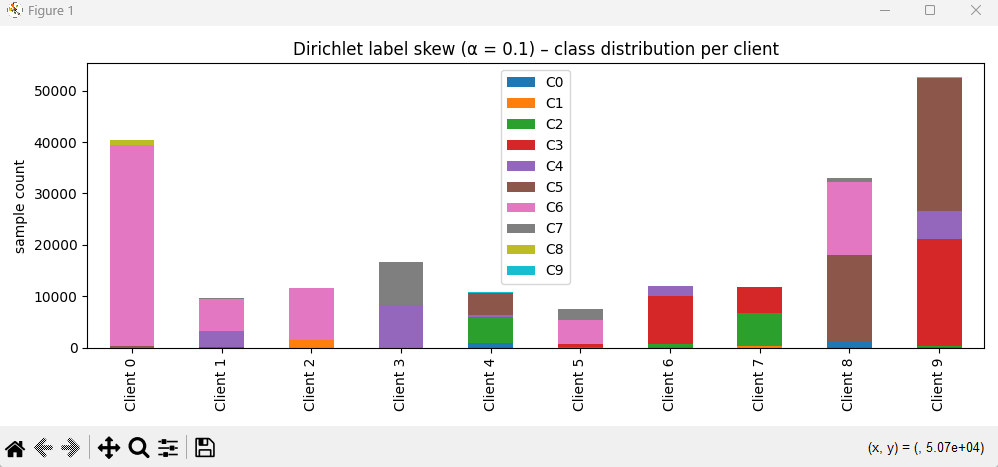
**Run 14:**

**FRACTION FIT = 0.5**

**FRACTION EVAL = 1**

**10 CLIENTS**

Non-IID alpha = 0.1



Duration: 108.74 seconds  
Rounds: 50

**Overall Performance (Round 50)**  
• Global Accuracy: 81.56%  
• Global F1 Score (macro): 0.4430  
• Global Precision (macro): 0.5689  
• Global Recall (macro): 0.4419

**Global Loss Trends**  
• Distributed Training Loss: 1.27 → 0.47 (round 1→50)  
• Centralized Evaluation Loss: 2.25 → 0.47

**Per Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.202 | 0.607 | 0.121 | Low recall despite decent precision; detection remains weak |
| 1 | 0.000 | 0.000 | 0.000 | Completely undetected across all rounds |
| 2 | 0.429 | 0.326 | 0.628 | Moderate; better recall than precision |
| 3 | 0.629 | 0.616 | 0.643 | Stable and balanced detection |
| 4 | 0.790 | 0.770 | 0.812 | Excellent performance; high precision and recall |
| 5 | 0.975 | 0.976 | 0.974 | Near-perfect classification |
| 6 | 1.000 | 0.999 | 1.000 | Perfect |
| 7 | 0.114 | 0.620 | 0.063 | Sparse detection; significant underperformance |
| 8 | 0.010 | 0.024 | 0.007 | Virtually undetected |
| 9 | 0.279 | 0.750 | 0.171 | Precision high, recall very weak |

**Run 15:**

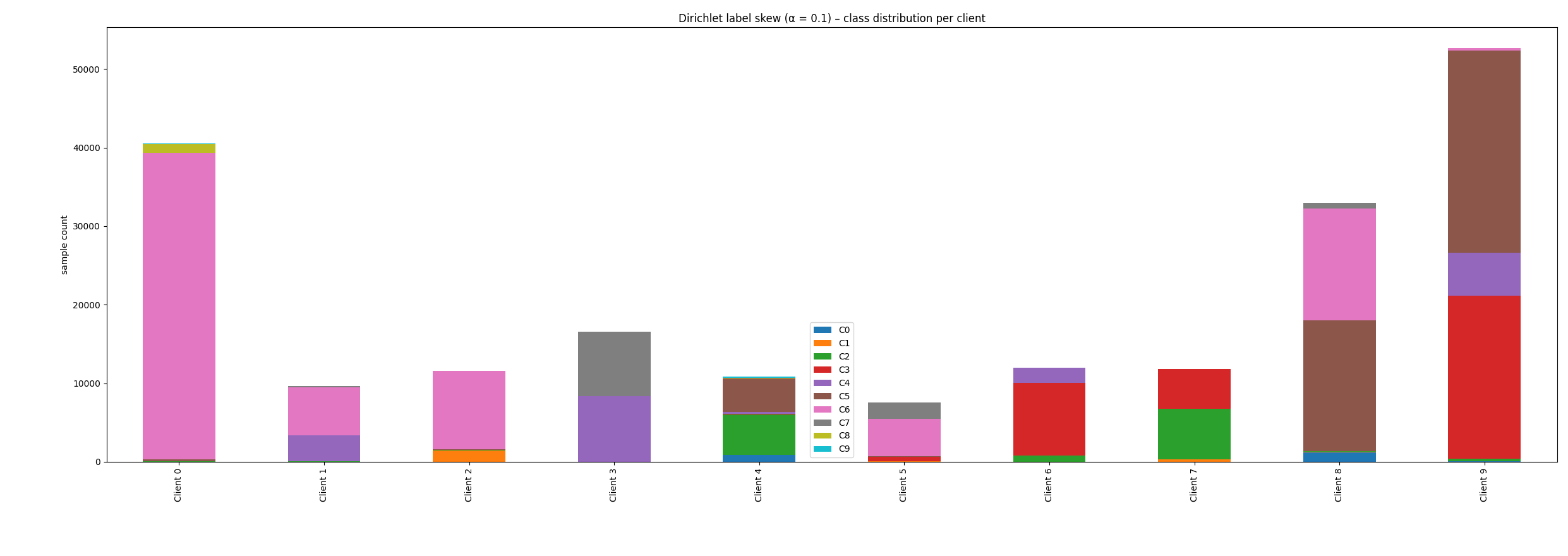
**FRACTION FIT = 0.5**

**FRACTION EVAL = 1**

**10 CLIENTS**

Non-IID alpha = 0.1

60 rounds



**Duration:** 117.97 seconds  
**Rounds:** 60

**Overall Performance (Round 60)**  
• Global Accuracy: 81.69%  
• Global F1 Score (macro): 0.4001  
• Global Precision (macro): 0.4806  
• Global Recall (macro): 0.4285

**Global Loss Trends**  
• Distributed Training Loss: 0.565 (final round)  
• Centralized Evaluation Loss: — *(not provided for each round)*

**Per Class Metrics at Round 60**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.089 | 0.862 | 0.047 | Severe recall bottleneck; misclassification persists |
| 1 | 0.000 | 0.000 | 0.000 | No detection across training |
| 2 | 0.121 | 0.337 | 0.074 | Minor improvement; still unreliable |
| 3 | 0.668 | 0.534 | 0.891 | High recall offsetting lower precision |
| 4 | 0.782 | 0.796 | 0.768 | Consistent and strong across all metrics |
| 5 | 0.974 | 0.973 | 0.974 | Robust, stable classification |
| 6 | 1.000 | 0.9997 | 0.9999 | Near-perfect as expected |
| 7 | 0.000 | 0.000 | 0.000 | Completely undetected |
| 8 | 0.175 | 0.119 | 0.331 | Recall up, but extremely low precision |
| 9 | 0.192 | 0.184 | 0.200 | Slightly balanced, but weak overall |

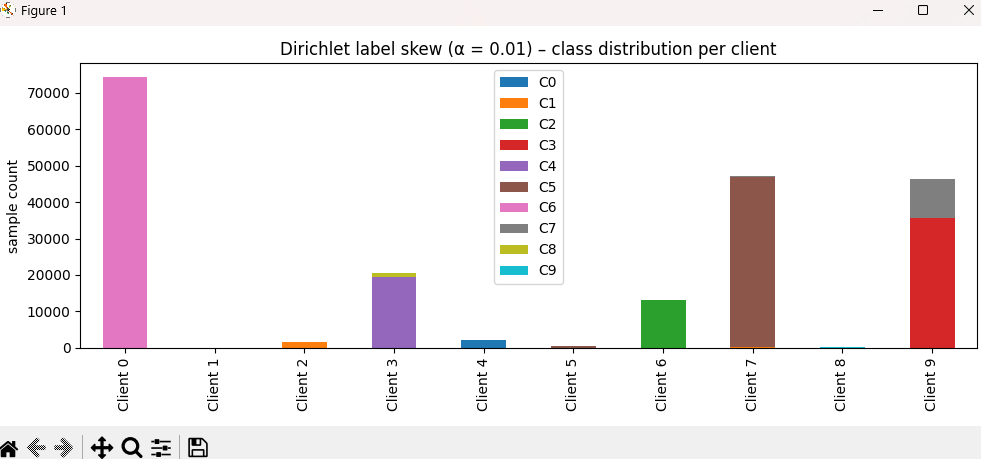
**Run 16:**

**FRACTION FIT = 0.5**

**FRACTION EVAL = 1**

**10 CLIENTS**

Non-IID alpha = 0.01



**Duration:** 102.18 seconds  
**Rounds:** 50

**Overall Performance (Round 50)**  
• Global Accuracy: 65.88%  
• Global F1 Score (macro): 0.2715  
• Global Precision (macro): 0.3536  
• Global Recall (macro): 0.3079  
→ *Highest Accuracy: 76.49% at Round 42*  
→ *Highest Macro F1 Score: 0.3308 at Round 42*

**Global Loss Trends**  
• Distributed Training Loss: 3.84 → 1.07 (round 1→50)  
• Centralized Evaluation Loss: 2.33 → 1.07  
→ *Lowest Evaluation Loss: 0.83 at Round 42*

**Per Class Metrics at Round 50**

| **Class** | **F1 Score** | **Precision** | **Recall** | **Notes** |
| --- | --- | --- | --- | --- |
| 0 | 0.000 | 0.000 | 0.000 | Completely undetected across all rounds |
| 1 | 0.013 | 0.030 | 0.009 | Very weak performance, slight recovery near end |
| 2 | 0.306 | 0.256 | 0.380 | Moderate recall, good late-stage improvement |
| 3 | 0.024 | 0.863 | 0.012 | High precision, extremely low recall |
| 4 | 0.451 | 0.355 | 0.618 | Solid recall, class improving over rounds |
| 5 | 0.755 | 0.642 | 0.915 | Strong and balanced classification |
| 6 | 0.953 | 0.909 | 1.000 | Near-perfect classification |
| 7 | 0.120 | 0.352 | 0.073 | Precision outweighs recall, minor class |
| 8 | 0.093 | 0.129 | 0.073 | Detected only in final round |
| 9 | 0.000 | 0.000 | 0.000 | Undetected throughout |